



TITLE:

ベイズ推測における事前分布の持つ情報の評価 (統計的モデルの新たな展望とそれに関連する話題)

AUTHOR(S):

小池, 健一; 神谷内, 裕

CITATION:

小池, 健一 ...[et al]. ベイズ推測における事前分布の持つ情報の評価 (統計的モデルの新たな展望とそれに関連する話題). 数理解析研究所講究録 2012, 1804: 1-10

ISSUE DATE:

2012-08

URL:

<http://hdl.handle.net/2433/194389>

RIGHT:

ベイズ推測における事前分布の持つ情報の評価

筑波大・数理物質系 小池 健一 (Ken-ichi Koike)

(Faculty of Pure and Applied Sciences, University of Tsukuba)

筑波大・教育研究科 神谷内 裕 (Yu Kamiyauchi)

(Master's Program in Education, University of Tsukuba)

1 はじめに

ベイズ推測は、推論方法の合理性や統計的決定理論への応用など、理論的に優れた性質を持っている。一方で、事前分布の取り方の恣意性や、事後分布の計算の煩雑さなど、ベイズ推測について様々な批判や議論がある。

ベイズ推測を用いた分析結果の正当性を判断する上で、仮定した事前分布が推測を行う前からどの程度の情報を持っているのか、すなわち、事前分布がどの程度の情報を追加して得られたものなのかを示すことは、重要な意味を持つといえる。したがって、事前分布の持つ情報の評価法について、統一的で一貫した理論や方法を提案することができれば、ベイズ推測を用いた分析結果の正当性を保証することができる。

ある推定精度を保証するのに必要な最小の標本数が、事前分布 π_1 の下では n_1 、事前分布 π_2 の下では n_2 (ただし $n_1 < n_2$) であったとする。このとき、事前分布 π_1 は π_2 よりも観測値 $n_2 - n_1$ 個分の情報を多く持っていると考えることができる。この考え方を推し進めることで、事前分布の持つ情報の大きさを、観測値の個数で表現することを試みる。

本稿では、事前分布の持つ情報の大きさを評価する方法について考察する。事前分布の持つ情報の評価法について、Clarke(1996)、森 (2010) などが提案している。本発表では、それぞれの方法について紹介し、それぞれの例や違い、問題点を考察する。またそれを踏まえ、事前分布の持つ情報の大きさを評価する方法を提案し、今後の研究課題について述べる。

2 先行研究における評価法

2.1 K-L 情報量による方法 (Clarke(1996))

Clarke(1996) は、事前分布 $\pi(\theta)$ が持つ情報を次のように定義した。無情報事前分布 $\pi_0(\theta)$ の下で標本数 n のデータ $X_n = (X_1, X_2, \dots, X_n)$ を観測し、事後分布が

$\pi_0(\theta|X_n)$ となったとする。このとき、 $\pi_0(\theta|X_n)$ と、情報の大きさを求めたい事前分布 $\pi(\theta)$ とのカルバック・ライブラー (K-L) 情報量

$$I(\pi_0(\theta|X_n), \pi(\theta)) = E_{\theta|X_n}[\log \pi_0(\theta|X_n)] - E_{\theta|X_n}[\log \pi(\theta)]$$

を考える。ただし $\theta|X_n$ は、 X_n を与えたときの θ の条件付確率分布に従う確率変数を表す。そして、 X_n, n を動かして、この $I(\pi_0(\theta|X_n), \pi(\theta))$ が最小になる観測値の個数 n を、この事前分布 $\pi(\theta)$ が持つ情報の大きさと定義した。ただし、ここでは無情報事前分布として、ジェフリーズ (Jeffreys) の事前分布を取るものとする。すなわち、 $\pi_0(\theta) \propto |I(\theta)|^{1/2}$ (ただし $I(\theta)$ はフィッシャー情報行列) である。

例 2.1 ポアソン分布

確率変数 X_i ($i = 1, 2, \dots, n$) をポアソン分布 $Po(\theta)$ に従う独立標本とする。ただし $\theta \in \Theta = [1, 11]$ に切断されているとする。パラメータ θ の事前分布 $\pi(\theta)$ は、正規分布 $N(6, 1)$ を $[1, 11]$ で切断した確率分布であると仮定する。このとき、事前分布 $\pi(\theta)$ の持つ情報の大きさを本方法で求める。

$\pi_0(\theta|X_n)$ を計算すると、 $\pi_0(\theta|X_n) = e^{-n\theta} \theta^{\sum x_i - \frac{1}{2}} / \int_1^{11} e^{-n\theta} \theta^{\sum x_i - \frac{1}{2}} d\theta$ となる。この式と、情報の大きさを求めたい事前分布 $\pi(\theta)$ との K-L 情報量

$$I(\pi_0(\theta|X_n), \pi(\theta)) = \int_1^{11} \pi_0(\theta|X_n) \log \frac{\pi_0(\theta|X_n)}{\pi(\theta)} d\theta$$

を、 $n = 1, 2, 3, \dots, \sum x_i = 1, 2, 3, \dots$ について計算し、それが最小となるときの n を求める。数値計算により、 $n = 6, \sum x_i = 37$ のとき、K-L 情報量は 0.01260359 で最小となる。したがって、事前分布 $\pi(\theta)$ として、正規分布 $N(6, 1)$ を $[1, 11]$ に切断した分布は、観測値 6 個分の情報を持っているということができる。

K-L 情報量による方法の問題点

例 2.1 において、K-L 情報量を $n = 1, 2, 3, \dots, \sum x_i = 1, 2, 3, \dots$ について計算し、それが最小となるような n を求めている。この計算方法では、 n や $\sum x_i$ は自然数で無限に存在するため、すべての n や $\sum x_i$ について計算することは不可能である。Clarke(1996) では、 $n = 6, \sum x_i = 37$ のとき、K-L 情報量は 0.01260359 で最小としている。極小性はこの計算方法で確かめられたとしても、最小性は確かめられない。

2.2 ベイズリスクによる方法 1(森 (2010))

森 (2010) は、事前分布 $\pi(\theta)$ の持つ情報の大きさを次のように定義した。

標本数 n の確率変数 X_i ($i = 1, 2, \dots, n$) について、パラメータ θ の事前分布を $\pi(\theta)$ 、 θ のベイズ推定量を δ_θ とし、 δ_θ のベイズリスクを $r_\pi(\delta_\theta) = E_\theta[E_{\mathbf{X}}[(\delta_\theta - \theta)^2 | \theta]]$ とする。また、標本数をさらに $n_{(+)}$ 個追加した確率変数 X_i ($i = 1, 2, \dots, n, n+1, \dots, n+n_{(+)}$) について、標本平均を \bar{X}' とし、 \bar{X}' のベイズリスクを $r_\pi(\bar{X}') = E_\theta[E_{\mathbf{X}}[(\bar{X}' - \theta)^2 | \theta]]$ とする。このとき、2つのベイズリスクが等しくなる、すなわち

$$r_\pi(\delta_\theta) = r_\pi(\bar{X}')$$

となるような追加標本の個数 $n_{(+)}$ を、事前分布 $\pi(\theta)$ の持つ情報の大きさと定義した。

例 2.2 ポアソン分布

確率変数 X_i ($i = 1, 2, \dots, n$) をポアソン分布 $Po(\lambda)$ に従う独立標本とし、パラメータ λ の事前分布としてガンマ分布 $Ga(\nu, \alpha)$ を仮定する。二乗誤差を損失関数とし、標本平均を \bar{X} とするとき、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。

λ のベイズ推定量は $\delta_\lambda = \frac{\nu + n\bar{X}}{\frac{1}{\alpha} + n}$ となることから、ベイズリスクを求めると、 $r_\pi(\delta_\lambda) = \frac{\nu\alpha}{\frac{1}{\alpha} + n}$ となる。一方、 $n_{(+)}$ 個の観測値を追加したときの標本平均からベイズリスクを求めると、 $r_\pi(\bar{X}') = \frac{\nu\alpha}{n + n_{(+)}}$ となる。2つのベイズリスクを等号でつなぎ、 $n_{(+)} = 1/\alpha$ を得る。したがって事前分布 $\pi(\lambda) \sim Ga(\nu, \alpha)$ は、 $1/\alpha$ 個分の情報を持っているといえることができる。

ベイズリスクによる方法 1 の問題点

比較相手となる推定量として、ベイズ推測とは無関係の標本平均 \bar{X}' を用いているが、標本平均を比較することがどのような状況でも妥当なのかといった根本的な疑問が起こる。例えば、正規分布 $N(\mu, \sigma^2)$ について、平均 μ の事前分布が持つ情報を推定するのに、事後平均 δ_μ の比較相手として標本平均 \bar{X}' を用いるのは妥当と考えられる。しかし、例えば分散 σ^2 の事前分布が持つ情報を推定する際にも、事後平均 δ_{σ^2} の比較相手として、同じように標本平均 \bar{X}' を用いるのは妥当とは言えない。

これについて森 (2010) では、不偏推定量や最尤推定量など、標本のみに基づいた標準的な推定方法が確立しているので、特定の推定量を採用するのではなく、状況に応じて推定量を変えればよい、としている。しかし、それは推定をする人がどの推定量を比較相手として選ぶかによって、得られる結論が変わってしまい、情報の大きさが一意に決まらないといった問題が生じる。

また、追加標本 $n_{(+)}$ が一般には標本数 n に依存してしまう。標本を観測する前から本来持っている事前分布の情報の大きさを定義する上で、標本を観測しなければわ

からないはずの標本数 n に依存してしまうことは、概念的に矛盾し、不都合である。

3 本研究で検討した評価法

- ベイズリスクによる方法 2
ベイズリスクによる方法 1 を基にして、比較する相手を標本平均のベイズリスクではなく、無情報事前分布から求めたベイズ推定量のベイズリスクを用いる
→ ベイズリスクによる方法 1 の問題点（比較相手の妥当性と評価方法の一般性）を改善するため
- 事後リスクによる方法
ベイズリスクによる方法 2 を基にして、両者のベイズリスクではなく、両者の事後リスクが等しくなる $n_{(+)}$ を求める
→ ベイズリスクによる方法 1 および 2 における、ベイズリスクの計算の煩雑さを改善するため
- 事後平均の漸近事後リスクによる方法
事後リスクによる方法を基にして、追加標本の考えをやめ、Hodges and Lehmann(1970) による漸近欠損量の概念を用いる
→ 事後リスクの計算の困難を改善し、ベイズリスクによる方法 1 以降で用いていた追加標本 $n_{(+)}$ の考えから生じる困難を改善するため
- 事後モードの漸近事後リスクによる方法
事後平均の漸近事後リスクによる方法を基にして、事後平均からの事後リスクではなく、事後モードからの事後リスクで比較する
→ 古典的統計学における最尤法の考えにつながり、最尤法における理論が活用できることが期待されるため
- 事後モードの事後リスクの漸近展開による方法
事後モードの漸近事後リスクによる方法と本質的には同じであるが、その計算過程で、Miyata(2008) の定理を用いる
- MLE の事後リスクの漸近展開による方法
前述の事後平均や事後モードからの事後リスクで比較するのではなく、Johnson(1970) の定理を用いて、MLE の事後リスクで比較する

3.1 ベイズリスクによる方法 2

ベイズリスクによる方法 1 の問題点（比較相手の妥当性と評価方法の一般性）を改善するため、比較する相手を標本平均のベイズリスクではなく、無情報事前分布から求めたベイズ推定量のベイズリスクを用いる。

例 3.1 ポアソン分布

設定は例 2.2 と同じとし、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。例 2.2 の結果から、 $r_\pi(\delta_\lambda) = \frac{\nu\alpha}{\frac{1}{\alpha}+n}$ となる。一方、標本数を $n_{(+)}$ 個追加したときのベイズ推定量（事後平均）からベイズリスクを求めると、ベイズリスクは $r_\pi(\hat{\delta}_\lambda) = \int_0^\infty \pi_0(\lambda) \left(\lambda + \frac{1}{4(n+n_{(+)})^2} \right) d\lambda$ で表されるが、 $\pi_0(\lambda) \propto 1/\sqrt{\lambda}$ であったため、積分が発散し、ベイズリスクを求められない。

ポアソン分布の例のように、事前分布が非正則な分布となった場合、この評価法では比較以前の問題が生じ、事前分布の持つ情報の大きさを決めることができない。

3.2 事後リスクによる方法

前節の方法の問題点（非正則な事前分布に対応できない問題）を改善するため、ベイズリスクではなく、事後リスク $E_\theta[(\delta_\theta - \theta)^2 | \mathbf{X}_n]$, $E_\theta[(\hat{\delta}_\theta - \theta)^2 | \mathbf{X}_{n+n_{(+)}}]$ が等しくなる $n_{(+)}$ を求める。本方法では二乗誤差を損失関数とするので、事後リスクは事後分布の分散に等しい。そのため事後分布がなんらかの確率分布に従う場合、事後リスクであれば計算が容易にできる。

例 3.2 ポアソン分布

設定は例 2.2 と同じとし、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。

λ の事後分布は、 $\pi(\lambda | \mathbf{X}_n) \sim Ga\left(\nu + n\bar{X}, \frac{1}{\frac{1}{\alpha} + n}\right)$ となる。ベイズ推定量（事後平均）の事後リスクは事後分布の分散に等しいので、 $E_\lambda[(\delta_\lambda - \lambda)^2 | \mathbf{X}_n] = \frac{\nu + n\bar{X}}{(\frac{1}{\alpha} + n)^2}$ となる。一方、標本数を $n_{(+)}$ 個追加したとき、ベイズ推定量（事後平均）の事後リスクは $E_\lambda[(\hat{\delta}_\lambda - \lambda)^2 | \mathbf{X}_{n+n_{(+)}}] = \frac{\bar{X}'}{n+n_{(+)}} + \frac{1}{2(n+n_{(+)})^2}$ となる。

2つの事後リスクを等号でつなぎ、 $n_{(+)}$ について解くことになるが、この2次方程式は計算が極めて困難であるうえ、解を求められたとしても複雑なものになってしまう。また、 $n_{(+)}$ は n や \bar{X} , \bar{X}' に依存した形となり、概念的に矛盾した結果になってしまう。

このように、追加標本 $n_{(+)}$ の考えを用いて事前分布の持つ情報の大きさを決める方法では、どうしても n に依存した形となってしまうため、この方法にも課題が残る。

3.3 事後平均の漸近事後リスクによる方法

事後リスクの計算の困難を改善し、ベイズリスクによる方法 1 以降で用いてきた追加標本 $n_{(+)}$ の考えから生じる困難を改善するため、Hodges and Lehmann(1970) および Lehmann(1983) で紹介されている漸近相対効率および漸近欠損量の概念、およびリスク関数の漸近展開の式を用いて、漸近事後リスクを求める。

前節で用いたそれぞれの事後リスクを $R_{1n} = E_\theta[(\delta_\theta - \theta)^2 | X_n]$, $R_{2n} = E_\theta[(\hat{\delta}_\theta - \theta)^2 | X_n]$ とする。ただし追加標本 $n_{(+)}$ の考えを用いないので、事後分布を求める過程で与える標本数はどちらも n で等しい。このとき、それぞれの事後リスク R_{1n}, R_{2n} を漸近展開すると、第 1 項 ($1/n$ の項) の係数は一致する。さらに $1/n^2$ の項まで展開して、

$$R_{1n} = \frac{a}{n} + \frac{b_1}{n^2} + o\left(\frac{1}{n^2}\right), \quad R_{2n} = \frac{a}{n} + \frac{b_2}{n^2} + o\left(\frac{1}{n^2}\right)$$

と表し (ただし $a > 0$)、2 つのモデルの標本数の差 d を

$$d = \frac{b_2 - b_1}{a} \quad (3.1)$$

とし (漸近欠損量)、これを元の事前分布の持つ情報の大きさと定義する。

例 3.3 ポアソン分布

設定は例 2.2 と同じとし、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。

例 3.2 で求めた 2 つの事後リスクを漸近展開する。ただしここでは、追加標本の考えを用いないので、 $n + n_{(+)}$ を n に直すことに注意。

$$R_{1n} = \frac{\bar{X}}{n} + \left(\nu - \frac{2\bar{X}}{\alpha}\right) \frac{1}{n^2} + o\left(\frac{1}{n^2}\right), \quad R_{2n} = \frac{\bar{X}}{n} + \frac{1}{2n^2}$$

これら 2 つの事後リスク R_{1n}, R_{2n} から標本数の差 d を求めると、 $d = \frac{1}{\bar{X}} \left(\frac{1}{2} - \nu\right) + \frac{2}{\alpha}$ となり、 d が \bar{X} に依存する問題が生じる。 \bar{X} に依存しない d を求めるため、ここでは d を \bar{X} の関数とみなし、 d の上界をとり、事前分布の持つ情報の大きさを定義する。

$\bar{X} > 0$, $\nu, \alpha > 0$ より、事前分布 $Ga(\nu, \alpha)$ は、

$$\sup d = \begin{cases} \frac{2}{\alpha} & (\nu \geq \frac{1}{2}), \\ \infty & (0 < \nu < \frac{1}{2}) \end{cases}$$

の大きさの情報を持っているということが出来る。

今後も上の例のように、 d が \bar{X} に依存する問題が生じる際は、その上界をとって情報の大きさを定義することにする。

3.4 事後モードの漸近事後リスクによる方法

事後平均の漸近事後リスクによる方法と同様の手順で、ここでは事後平均ではなく事後モードの事後リスクを求めて比較する。

事後モードの事後リスクをそれぞれ R_{1n}^*, R_{2n}^* とし、漸近展開すると、第1項 ($1/n$ の項) の係数は一致する。さらに $1/n^2$ の項まで展開して、

$$R_{1n}^* = \frac{a}{n} + \frac{b_1}{n^2} + o\left(\frac{1}{n^2}\right), \quad R_{2n}^* = \frac{a}{n} + \frac{b_2}{n^2} + o\left(\frac{1}{n^2}\right)$$

と表し (ただし $a > 0$)、前節 (3.1) 式の漸近欠損量を用いて2つのモデルの標本数の差 d を求め、事前分布の持つ情報の大きさを定義する。

無情報事前分布がルベーク測度に比例したもの (もしくは一様分布) になるとき、事後モードは、古典的統計学における最尤法と全く同じになる。したがってこの方法を用いることで、古典的統計学における最尤法のさまざまな理論を活用できることが期待される。

また、 R_{1n}^* や R_{2n}^* は、前節の事後平均の漸近事後リスクによる方法で求めた R_{1n} や R_{2n} 、 δ_θ 、 $\hat{\delta}_\theta$ を用いて

$$R_{1n}^* = R_{1n} + (\delta_\theta - \delta_\theta^*)^2, \quad R_{2n}^* = R_{2n} + (\hat{\delta}_\theta - \hat{\delta}_\theta^*)^2$$

と計算できる。

例 3.4 ポアソン分布

設定は例 2.2 と同じとし、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。

それぞれの事後モードを求め、そこから計算される2つの事後リスクを漸近展開すると、

$$R_{1n}^* = \frac{\bar{X}}{n} + \left(\nu + 1 - \frac{2\bar{X}}{\alpha}\right) \frac{1}{n^2} + o\left(\frac{1}{n^2}\right), \quad R_{2n}^* = \frac{\bar{X}}{n} + \frac{3}{2n^2}$$

となる。これら2つの事後リスク R_{1n}^*, R_{2n}^* から標本数の差 d を求めると、 $d = \frac{1}{\bar{X}} \left(\frac{1}{2} - \nu\right) + \frac{2}{\alpha}$ となり、前述の例 3.3 で求めた d と全く同じ式になる。したがって例 3.3 と同じ結論が得られ、事前分布 $Ga(\nu, \alpha)$ は、

$$\begin{cases} \frac{2}{\alpha} & (\nu \geq \frac{1}{2}), \\ \infty & (0 < \nu < \frac{1}{2}) \end{cases}$$

の大きさの情報を持っているといえることができる。

3.5 事後モードの事後リスクの漸近展開による方法

前節と同様に、事後モードの事後リスク $E_\theta[(\delta_\theta^* - \theta)^2 | \mathbf{X}_n]$ を計算するが、事後リスクの漸近展開に関する定理 (Miyata(2008)) を用いる。そこから漸近欠損量を用いて標本数の差 d を求め、事前分布の持つ情報の大きさを定義する。

$h_n(\theta) = -n^{-1} \log f(\mathbf{X}_n | \theta) \pi(\theta)$ とし、 δ_θ^* を事後モードとする。また、上付きの (k) や l は導関数を表すものとする。以下、記号 (θ) には $\theta = \delta_\theta^*$ を代入したものとし、省略して表記する。例えば、 h_n'' は $h_n''(\delta_\theta^*)$ を表す。さらに、 $\sigma_n = 1/\sqrt{nh_n''}$ とする。

定理 3.1 (Miyata(2008)) 事後リスク $E_\theta[(\theta - \delta_\theta^*)^2 | \mathbf{X}_n]$ は以下のように漸近展開できる。

$$E_\theta[(\theta - \delta_\theta^*)^2 | \mathbf{X}_n] = \sigma_n^2 + \frac{\sigma_n^2}{n} \left(-\frac{h_n^{(4)}}{2(h_n'')^2} + \frac{5(h_n^{(3)})^2}{4(h_n'')^3} \right) + o\left(\frac{1}{n^2}\right)$$

本節と前節は計算方法の違いのみであり、本節で得られる結論はすべて前節と同じである。

定理 3.1 を用いることで、正規分布やベルヌーイ分布、ポアソン分布などの限られた分布だけでなく、より一般的な分布についても情報の大きさを決められることが期待できる。

3.6 MLE の事後リスクの漸近展開による方法

事後平均の事後リスクや、事後モードの事後リスクを用いるのではなく、本節では新しく最尤推定量 (MLE) の事後リスクを用いて比較する方法についても考察する。MLE の事後リスクの計算は Johnson(1970) の定理を用いる。

定理 3.2 (Johnson(1970)) $\bar{h}_n(\theta) = -n^{-1} \log f(\mathbf{X}_n | \theta)$ とし、 $\bar{\delta}_\theta$ を θ の最尤推定量 (MLE) とする。また、上付きの (k) や l は導関数を表すものとする。以下、記号 (θ) には $\theta = \bar{\delta}_\theta$ を代入したものとし、省略して表記する。例えば、 \bar{h}_n'' は $\bar{h}_n''(\bar{\delta}_\theta)$ を表し、 π は $\pi(\bar{\delta}_\theta)$ を表す。さらに、 $\bar{\sigma}_n^2 = 1/n\bar{h}_n''$ とする。

このとき、事後リスク $E_\theta[(\theta - \bar{\delta}_\theta)^2 | \mathbf{X}_n]$ の漸近展開が以下の式で計算できる。

$$E_\theta[(\theta - \bar{\delta}_\theta)^2 | \mathbf{X}_n] = \bar{\sigma}_n^2 + \frac{\bar{\sigma}_n^2}{n} \left(-\frac{\bar{h}_n^{(4)}}{2(\bar{h}_n'')^2} + \frac{5(\bar{h}_n^{(3)})^2}{4(\bar{h}_n'')^3} - \frac{2\bar{h}_n^{(3)}\pi'}{(\bar{h}_n'')^2\pi} + \frac{\pi''}{\bar{h}_n''\pi} \right) + o\left(\frac{1}{n^2}\right)$$

例 3.5 ポアソン分布

設定は例 2.2 と同じとし、事前分布 $\pi(\lambda)$ の持つ情報の大きさを本方法で求める。計算結果のみ記す。2つの事後リスクから標本数の差 d を求めると、

$$d = \frac{-4\bar{X}^2 + 8\alpha\bar{X}(1 + \nu) + \alpha^2(3 - 4\nu - 4\nu^2)}{4\alpha^2\bar{X}}$$

となる。 d の上界をとると、事前分布 $Ga(\nu, \alpha)$ は、

$$\begin{cases} \frac{2(\nu+1) - \sqrt{4\nu^2 + 4\nu - 3}}{\alpha} & (\nu \geq \frac{1}{2}), \\ \infty & (0 < \nu < \frac{1}{2}) \end{cases}$$

の大きさの情報を持っているといえることができる。

4 今後の課題

事前分布の持つ情報の評価について、森 (2010) によるベイズリスクによる方法 1 を除き、本発表で用いたいずれの方法についても、比較相手となる無情報事前分布をどのように設定すべきかという問題がある。本発表では一貫してジェフリーズの事前分布を比較相手として用いてきたが、前述の通り、あらゆる事前分布の中で最も情報が小さい事前分布というわけではない。また、そもそも情報の大きさ自体をどのように定義するかといったことについても議論の必要がある。あらゆる事前分布の中で最も情報が小さい事前分布を設定する方法が確立できれば、興味の対象となる事前分布の持つ情報の大きさについて、常に非負の値で定義することができる。

また、事前分布の持つ情報の評価法について、特定の分布（統計モデル）のみならず、より広く一般的な場合（統計モデル）についても適用可能で、かつ well-defined な定義が必要とされる。well-defined な定義とは、事前分布の持つ情報の大きさが標本数 n や推定量に依存しないこと、評価法が特定の推定量によらずに確立できることを指す。標本を観測する前に持っている事前分布の情報の大きさが、標本を観測しなければわからないはずの標本数 n や推定量に依存するのでは、概念的に矛盾する。また特定の推定量によって評価方法が変わってしまうのでは、そのときの方法によって得られる結論が変わってしまうため、well-defined な定義とはならない。

さらに、計算の容易さについても望ましい性質の一つである。これらを追求することが今後の研究課題である。

参考文献

- [1] 安道知寛. ベイズ統計モデリング. 朝倉書店, 2010.

- [2] Bernardo, J. M. Reference posterior distributions for Bayesian Inference. *J. R. Statist. Soc. B*, (1979), **41**, 113–147.
- [3] Clarke, B. Implications of reference priors for prior information and for sample size. *J. Amer. Statist. Assoc.*, (1996), **91**, 173–184.
- [4] Ghosh, J. K. *et al. An Introduction to Bayesian Analysis: Theory and Methods*. Springer, 2006.
- [5] Hodges, J. K. and Lehmann, E. L. Deficiency. *Ann. Math. Statist.*, (1970), **41**, 783–801.
- [6] Johnson, R. A. An asymptotic expansions for posterior distributions. *Ann. Math. Statist.*, (1967), **38**, 1899–1907.
- [7] Johnson, R. A. Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.*, (1970), **41**, 851–864.
- [8] 小西貞則・北川源四郎. 情報量基準. 朝倉書店, 2004.
- [9] Lehmann, E. L. *Theory of Point Estimation*. Wiley, 1983.
- [10] 松原望. ベイズ統計学概説：フィッシャーからベイズへ. 培風館, 2010.
- [11] Miyata, Y. Higher order expansions for posterior distributions using posterior modes. *J. Japan Statist. Soc.*, (2008) **38**(3), 415–429.
- [12] 森治憲. ベイズ法における事前分布の持つ情報の評価法 日本統計学会誌. (2010), **40**(1), 1–22.
- [13] 中妻照雄. 入門ベイズ統計学. 朝倉書店, 2007.
- [14] Press, S. J. *et al. Subjective and objective Bayesian statistics: principles, models, and applications, 2nd ed.* Wiley-Interscience, 2003.